WHAT IS CLAIMED IS:

1. An apparatus for producing an acoustic model for speech recognition, said apparatus comprising:

means for categorizing a plurality of first noise samples into clusters, a number of said clusters being smaller than that of noise samples;

means for selecting a noise sample in each of the clusters to set the selected noise samples to second noise samples for training;

means for storing thereon an untrained acoustic model for training; and

means for training the untrained acoustic model by using the second noise samples for training so as to produce the acoustic model for speech recognition.

2. An apparatus for producing an acoustic model according to claim 1, wherein said categorizing means further comprises:

means for executing a speech analysis of each of the first noise samples by frame to obtain characteristic parameters for each frame in each of the first noise samples;

means for obtaining a time-average vector in each of the characteristic vectors of each of the first noise samples; and

means for clustering the time-average vectors of the respective characteristic vectors into the clusters.

3. An apparatus for producing an acoustic model according to claim 2, wherein said clustering means performs the clustering operation

by using a hierarchical clustering method.

     4. An apparatus for producing an acoustic model according to claim 2, wherein said clustering means further comprises:

5
     means for setting the time-average vectors to clusters;

     means for computing each distance between each cluster;

     means for extracting at least one pair of two clusters in the set clusters, said at least one pair of clusters providing a distance which is a shortest in other any paired two clusters in the set clusters;

10
     means for linking the two extracted clusters to set the linked clusters to a same cluster;

     means for determining whether or not a number of the clusters including the same cluster equals to one, said extracting means and linking means performing the extracting operation and the linking operation

15
repeatedly in a case where the determination is that the number of clusters does not equal to one;

     means for producing, in a case where the determination is that the number of clusters equals to one, a dendrogram indicating a linking relationship between the linked clusters and indicating similarities among

20
the first noise samples; and

     means for cutting the dendrogram at a predetermined position thereof to obtain plural clusters linked to each other,

     wherein said selecting means selects the noise sample in each of the obtained plural clusters.

25

     5. An apparatus for producing an acoustic model according to

claim 1, wherein said training means further comprises:

means for storing thereon a plurality of speech samples for training;

means for extracting at least one of the second noise samples for training;

means for superimposing the at least one of extracted second noise sample on the speech samples for training;

means for executing a speech analysis of each of the noise superimposed speech samples by frame to obtain characteristic parameters corresponding to the noise superimposed speech samples; and

means for training the untrained acoustic model on the basis of the obtained characteristic parameters to obtain the acoustic model for speech recognition, said trained acoustic model being trained according to the at least one extracted noise sample.

6. An apparatus for recognizing an unknown speech signal comprising:

means for categorizing a plurality of first noise samples into clusters, a number of said clusters being smaller than that of noise samples;

means for selecting a noise sample in each of the clusters to set the selected noise samples to second noise samples for training;

means for storing thereon an untrained acoustic model for training;

means for training the untrained acoustic model by using the second noise samples for training so as to obtain a trained acoustic model for speech recognition;

means for inputting the unknown speech signal; and

means for recognizing the unknown speech signal on the basis of

the trained acoustic model for speech recognition.

7. A programmed-computer readable storage medium comprising:

means for causing a computer to categorize a plurality of first noise samples into clusters, a number of said clusters being smaller than that of noise samples;

means for causing a computer to select a noise sample in each of the clusters to set the selected noise samples to second noise samples for training;

means for causing a computer to store thereon an untrained acoustic model; and

means for causing a computer to train the untrained acoustic model by using the second noise samples for training so as to produce an acoustic model for speech recognition.

8. A method of producing an acoustic model for speech recognition, said method comprising the steps of:

preparing a plurality of first noise samples;

preparing an untrained acoustic model for training;

categorizing the plurality of first noise samples into clusters, a number of said clusters being smaller than that of noise samples;

selecting a noise sample in each of the clusters to set the selected noise samples to second noise samples for training; and

training the untrained acoustic model by using the second noise samples for training so as to produce the acoustic model for speech recognition.